

Optimal order quadrature error bounds for infinite-dimensional higher order digital sequences

Takashi Goda*, Kosuke Suzuki†, Takehito Yoshiki‡

November 22, 2016

Abstract

Quasi-Monte Carlo (QMC) quadrature rules using higher order digital nets and sequences have been shown to achieve the almost optimal rate of convergence of the worst-case error in Sobolev spaces of arbitrary fixed smoothness $\alpha \in \mathbb{N}$, $\alpha \geq 2$. In a recent paper by the authors, it was proved that randomly-digitally-shifted order 2α digital nets in prime base b achieve the best possible rate of convergence of the root mean square worst-case error of order $N^{-\alpha}(\log N)^{(s-1)/2}$ for $N = b^m$, where N and s denote the number of points and the dimension, respectively, which implies the existence of an optimal order QMC rule. More recently, the authors provided an explicit construction of such an optimal order QMC rule by using Chen-Skriganov's digital nets in conjunction with Dick's digit interlacing composition. These results were for fixed number of points. In this paper we give a more general result on an explicit construction of optimal order QMC rules for arbitrary fixed smoothness $\alpha \in \mathbb{N}$ including the endpoint case $\alpha = 1$. That is, we prove that the projection of any infinite-dimensional order $2\alpha + 1$ digital sequence in prime base b onto the first s coordinates achieves the best possible rate of convergence of the worst-case error of order $N^{-\alpha}(\log N)^{(s-1)/2}$ for $N = b^m$. The explicit construction presented in this paper is not only easy to implement but also extensible in both N and s .

Keywords: Quasi-Monte Carlo, Numerical integration, Higher order digital sequences, Sobolev space

MSC classifications: Primary, 41A55, 65D32; Secondary, 42C10, 65C05, 65D30

1 Introduction and statement of the main result

In this paper we study numerical integration of smooth functions defined over the s -dimensional unit cube. Let $f: [0, 1]^s \rightarrow \mathbb{R}$ be an integrable function. We denote the true integral of f by

$$I(f) := \int_{[0,1]^s} f(\mathbf{x}) \, d\mathbf{x}.$$

*Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan (goda@frcer.t.u-tokyo.ac.jp)

†School of Mathematics and Statistics, The University of New South Wales, Sydney 2052, Australia. (kosuke.suzuki1@unsw.edu.au)

‡School of Mathematics and Statistics, The University of New South Wales, Sydney 2052, Australia. (takehito.yoshiki1@unsw.edu.au)

A quasi-Monte Carlo (QMC) rule is an equal-weight quadrature rule where the weights sum up to 1. That is, by using an N -element point set $P_N^{(s)} = \{\mathbf{x}_n^{(s)} : 0 \leq n < N\} \subset [0, 1]^s$, a QMC rule approximates $I(f)$ by

$$I(f; P_N^{(s)}) = \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{x}_n^{(s)}).$$

In case of an infinite sequence of points $\mathcal{S}^{(s)} = \{\mathbf{x}_n^{(s)} : n \geq 0\} \subset [0, 1]^s$, we use the first N elements of $\mathcal{S}^{(s)}$ as $P_N^{(s)}$. Moreover, in case of an infinite-dimensional sequence of points $\mathcal{S} = \{\mathbf{x}_n : n \geq 0\} \subset [0, 1]^{\mathbb{N}}$, we use the first N points of the projection of \mathcal{S} on to the first s coordinates as $P_N^{(s)}$. An explicit construction of a good infinite-dimensional sequence of points is of considerable importance in practical applications since then the point set is extensible in both N and s , i.e., both the number of points and the dimension can be increased while retaining the existing points [20].

In order to measure the quality of a QMC point set or sequence for a class of integrands instead of a single integrand, we consider the so-called worst-case error for a normed function space which is defined as follows. Let V be a function space with norm $\|\cdot\|_V$. The worst-case error of a point set $P_N^{(s)}$ in V , denoted by $e^{\text{wor}}(V, P_N^{(s)})$, is the supremum of the quadrature error in the unit ball of V , i.e.,

$$e^{\text{wor}}(V, P_N^{(s)}) := \sup_{\substack{f \in V \\ \|f\|_V \leq 1}} |I(f; P_N^{(s)}) - I(f)|.$$

Reproducing kernel Hilbert spaces of Sobolev type consisting of functions which have square integrable partial mixed derivatives up to order 1 in each variable have been frequently studied as an important example of a normed function space in the literature, see for instance [18] and [30, Appendix A]. The worst-case error in such function spaces is often connected with geometric discrepancy of point sets, which measures how uniformly distributed the point set is. This fact has motivated to construct low-discrepancy point sets and sequences. Indeed there are many known explicit constructions of (possibly infinite-dimensional) low-discrepancy point sets and sequences including those of Halton [17], Sobol' [34], Faure [13], Niederreiter [27], Niederreiter and Xing [29], Chen and Skrivogov [4] and Dick and Pillichshammer [12] to list just a few.

Recently, function spaces with higher smoothness $\alpha \in \mathbb{N}$, $\alpha \geq 2$, have attracted considerable attention in the area of uncertainty quantification, in particular partial differential equations with random coefficients, see for instance [9, 10]. For such smooth functions it is possible to achieve higher order of convergence of the worst-case error by exploiting the smoothness of the functions. In his seminal work [5, 6], Dick studied the decay of the Walsh coefficients of smooth functions and introduced a digit interlacing composition for constructing the so-called (possibly infinite-dimensional) higher order digital nets and sequences which achieve the almost optimal rate of convergence $N^{-\alpha}(\log N)^{c(s, \alpha)}$ for $N = b^m$ and some $c(s, \alpha) > 0$, see [6, Theorem 5.4 & Corollary 5.5] and also [2, Theorem 22]. (We refer to Subsection 2.1 for the definition of higher order digital nets and sequences.) Although this order of convergence is best possible up to some power of a $\log N$ factor, it has been unknown until recently whether the exponent $c(s, \alpha)$ can be improved to be optimal.

There has been some progress reported on the optimality of the exponent $c(s, \alpha)$ for higher order digital nets. In [22], Hinrichs et al. considered periodic Sobolev spaces and periodic Nikol'skij-Besov spaces with (real-valued) dominating mixed smoothness up to 2, and obtained $c(s, \alpha) = (s-1)/2$ for order 2 digital nets, which is best possible. In [15], the authors of this paper considered a reproducing kernel Hilbert space $H_{\alpha,s}$ of Sobolev type consisting of non-periodic functions with smoothness $\alpha \in \mathbb{N}$, $\alpha \geq 2$, and proved that randomly-digitally-shifted order 2α digital nets can achieve the rate of convergence of the root mean square worst-case error of order $N^{-\alpha}(\log N)^{(s-1)/2}$. (We refer to Subsection 2.2 for the definition of $H_{\alpha,s}$.) This result directly implies the existence of a QMC rule achieving $c(s, \alpha) = (s-1)/2$, which is best possible [15, Proposition 1.3]. More recently in [16], the authors of this paper provided an explicit construction of such an optimal order QMC rule by using Chen-Skriganov's digital nets in conjunction with Dick's digit interlacing composition.

In this third paper of the authors on the optimality of higher order digital nets and sequences, we prove the following theorem:

Theorem 1. *Let b be a prime, and $s, \alpha \in \mathbb{N}$. Let \mathcal{S} be an infinite-dimensional order $2\alpha + 1$ digital sequence over the finite field \mathbb{F}_b , and let $P_N^{(s)}$ be the set of the first N points of the projection of \mathcal{S} onto the first s coordinates for $N \in \mathbb{N}$. Then for any $m \in \mathbb{N}$, we have*

$$e^{\text{wor}}(H_{\alpha,s}; P_{b^m}^{(s)}) \leq C_{\alpha,b,s} \frac{m^{(s-1)/2}}{b^{\alpha m}},$$

where $C_{\alpha,b,s}$ is positive and independent of m .

Theorem 1 claims that the projection of any infinite-dimensional order $2\alpha + 1$ digital sequence onto the first s coordinates achieves the best possible rate of convergence $N^{-\alpha}(\log N)^{(s-1)/2}$ of the worst-case error in $H_{\alpha,s}$, when N is of the form b^m for some $m \in \mathbb{N}$ with a fixed prime base b . Thus our result implies a new explicit construction of optimal order QMC rules for smooth integrands, which is not only easy to implement but also extensible in both N and s . As shown in [21, 31], in order for an extensible QMC rule to achieve a rate of convergence better than N^{-1} , the number of points must grow at least geometrically. Therefore, our result is also best possible in this regard when $\alpha \geq 2$.

We would remark that our present result covers the endpoint case $\alpha = 1$, which has not been included in the previous results [15, 16]. Furthermore, since any order $2\alpha + 1$ digital sequence is also an order $2\alpha' + 1$ digital sequence for all $\alpha' < \alpha$ due to the so-called propagation rules [3], it also achieves the best possible rate of convergence $N^{-\alpha'}(\log N)^{(s-1)/2}$ of the worst-case error in $H_{\alpha',s}$ as long as $\alpha' < \alpha$. The main idea of our proof of Theorem 1 is to exploit the decay and the sparsity of the Walsh coefficients of the reproducing kernel simultaneously in a different way from [16]. A similar approach is used in [12] to prove the best possible order of the L_2 discrepancy bound for higher order digital sequences over \mathbb{F}_2 . Whether it is possible to lower the necessary order of digital sequences from $2\alpha + 1$ while still achieving the best possible rate of convergence is an open question.

We note that the optimal order of convergence in the Sobolev space with smoothness α is also achieved by using the Frolov lattice rule in conjunction

with a periodization technique [14, 25, 36, 37]. In fact, it has been shown that the Frolov lattice rule achieves the optimal order of convergence in Besov and Triebel-Lizorkin spaces with dominating mixed smoothness, and the optimality in the Sobolev space is obtained from the fact that the Sobolev space coincides with the Triebel-Lizorkin space under the special choice of the parameters. Although the Frolov lattice rule is an equal-weight quadrature rule, the weights do not sum up to 1 in general, so that it is not a QMC rule. Periodizing an integrand can be interpreted as having a modified quadrature rule with possibly unequal weights of the original integrand [19], and is shown to cause a severe consequence especially in high dimensions even for the constant functions [24]. Since our optimal order QMC rule does not require such a periodization, a numerical stability is kept independently of the dimension. Nevertheless, a significant advantage of the Frolov lattice rule lies in its universality, that is, one quadrature rule satisfies the optimal order worst-case error bounds for many function spaces (such as Besov, Triebel-Lizorkin, and Sobolev) with different smoothness. As mentioned above, our presented optimal order QMC rules also possess a kind of universality, which is though weaker than that of the Frolov lattice rule. On the other hand, the Frolov lattice rule is not extensible in either N or s and also seems not easy to implement in high dimensions. Finally we would point out that both of the quadrature rules have randomized and unbiased variants [8, 23].

2 Preliminaries

Throughout this paper we shall use the following notations. Let \mathbb{N} be the set of positive integers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Let b be a fixed prime, and \mathbb{F}_b the finite field with b elements, which we identify with the set $\{0, 1, \dots, b-1\}$ equipped with addition and multiplication modulo b . The operators \oplus and \ominus denote digitwise addition and subtraction modulo b , respectively, that is, for $k = \sum_{i=1}^{\infty} \kappa_i b^{i-1} \in \mathbb{N}_0$ and $k' = \sum_{i=1}^{\infty} \kappa'_i b^{i-1} \in \mathbb{N}_0$ with $\kappa_i, \kappa'_i \in \mathbb{F}_b$, where all, except a finite number of κ_i and κ'_i , are 0, we define

$$k \oplus k' := \sum_{i=1}^{\infty} \lambda_i b^{i-1} \quad \text{and} \quad k \ominus k' := \sum_{i=1}^{\infty} \lambda'_i b^{i-1},$$

where $\lambda_i = \kappa_i + \kappa'_i \pmod{b}$ and $\lambda'_i = \kappa_i - \kappa'_i \pmod{b}$. In case of vectors in \mathbb{N}_0^s , the operators \oplus and \ominus are applied componentwise.

2.1 Higher order digital nets and sequences

2.1.1 Definitions

We start with a general digital construction scheme of infinite-dimensional finite point sets in the unit cube due to Niederreiter [28, p. 63].

Definition 1. For $m, n \in \mathbb{N}$, let $C_j \in \mathbb{F}_b^{n \times m}$ for $j \in \mathbb{N}$ be $n \times m$ matrices over \mathbb{F}_b . For an integer $0 \leq h < b^m$, we denote the b -adic expansion of h by $h = \sum_{i=1}^m \eta_i b^{i-1}$. For $j \in \mathbb{N}$, let

$$x_{h,j} = \frac{\xi_{1,h,j}}{b} + \frac{\xi_{2,h,j}}{b^2} + \cdots + \frac{\xi_{n,h,j}}{b^n},$$

where $\xi_{1,h,j}, \xi_{2,h,j}, \dots, \xi_{n,h,j}$ are given by

$$(\xi_{1,h,j}, \xi_{2,h,j}, \dots, \xi_{n,h,j})^\top = C_j(\eta_1, \eta_2, \dots, \eta_m)^\top.$$

Then the set $P_{b^m} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{b^m-1}\}$ with $\mathbf{x}_h = (x_{h,1}, x_{h,2}, \dots) \in [0, 1]^\mathbb{N}$ is called an infinite-dimensional digital net over \mathbb{F}_b with generating matrices C_j . Moreover, the projection of P_{b^m} onto the first s coordinates is called a digital net over \mathbb{F}_b with generating matrices C_1, \dots, C_s .

This construction scheme can be extended to the case of an infinite-dimensional sequence of points [28, p. 72].

Definition 2. Let $C_j \in \mathbb{F}_b^{\mathbb{N} \times \mathbb{N}}$ for $j \in \mathbb{N}$ be $\mathbb{N} \times \mathbb{N}$ matrices over \mathbb{F}_b . For each $C_j = (c_{k,l}^{(j)})_{k,l \in \mathbb{N}}$ we assume that there exists a function $K : \mathbb{N} \rightarrow \mathbb{N}$ such that $c_{k,l}^{(j)} = 0$ when $k > K(l)$. For a non-negative integer h , we denote the b -adic expansion of h by $h = \sum_{i=1}^a \eta_i b^{i-1}$ for some $a \in \mathbb{N}$. For $j \in \mathbb{N}$, let

$$x_{h,j} = \frac{\xi_{1,h,j}}{b} + \frac{\xi_{2,h,j}}{b^2} + \dots,$$

where $\xi_{1,h,j}, \xi_{2,h,j}, \dots$ are given by

$$(\xi_{1,h,j}, \xi_{2,h,j}, \dots)^\top = C_j(\eta_1, \eta_2, \dots, \eta_a, 0, 0, \dots)^\top.$$

Then the sequence $\mathcal{S} = (\mathbf{x}_0, \mathbf{x}_1, \dots)$ with $\mathbf{x}_h = (x_{h,1}, x_{h,2}, \dots) \in [0, 1]^\mathbb{N}$ is called an infinite-dimensional digital sequence over \mathbb{F}_b with generating matrices C_j . Moreover, the projection of \mathcal{S} onto the first s coordinates is called a digital sequence over \mathbb{F}_b with generating matrices C_1, \dots, C_s .

Note that the condition $c_{k,l}^{(j)} = 0$ for all sufficiently large k is a standard assumption to ensure that every number $x_{h,j}$ is b -adic rational, i.e., $x_{h,j}$ is written in a finite b -adic expansion, so that every point in \mathcal{S} belongs to $[0, 1]^\mathbb{N}$, see for instance [12, 28].

Now we give the definitions of higher order digital nets and sequences, which are special cases of [6, Definition 4.3] and [6, Definition 4.8], respectively.

Definition 3. For $m, n, s, \alpha \in \mathbb{N}$ with $n \geq \alpha m$, let $P_{b^m}^{(s)}$ be a digital net over \mathbb{F}_b with generating matrices $C_1, \dots, C_s \in \mathbb{F}_b^{n \times m}$. For $1 \leq i \leq n$ and $1 \leq j \leq s$, we denote the i -th row of C_j by $\mathbf{c}_{i,j}$. Let t be an integer with $0 \leq t \leq \alpha m$ which satisfies the following condition: For all $1 \leq i(j, v_j) < \dots < i(j, 1) \leq n$ with

$$\sum_{j=1}^s \sum_{l=1}^{\min(\alpha, v_j)} i(j, l) \leq \alpha m - t,$$

the vectors $\mathbf{c}_{i(1, v_1), 1}, \dots, \mathbf{c}_{i(1, 1), 1}, \dots, \mathbf{c}_{i(s, v_s), s}, \dots, \mathbf{c}_{i(s, 1), s}$ are linearly independent over \mathbb{F}_b . Then we call $P_{b^m}^{(s)}$ an order α digital (t, m, s) -net over \mathbb{F}_b .

Definition 4. Let \mathcal{S} be an infinite-dimensional digital sequence over \mathbb{F}_b with generating matrices $C_j \in \mathbb{F}_b^{\mathbb{N} \times \mathbb{N}}$. For $s \in \mathbb{N}$, $\mathcal{S}^{(s)}$ denotes the projection of \mathcal{S} onto the first s coordinates, i.e., the digital sequence over \mathbb{F}_b with generating matrices C_1, \dots, C_s . Let $t(s)$ be a non-negative integer. If for any $m \in \mathbb{N}$ with $\alpha m \geq t(s)$

the set of the first b^m points of $\mathcal{S}^{(s)}$ is an order α digital $(t(s), m, s)$ -net over \mathbb{F}_b , then we call $\mathcal{S}^{(s)}$ an order α digital $(t(s), s)$ -sequence over \mathbb{F}_b .

Moreover, if there exists a function $t: \mathbb{N} \rightarrow \mathbb{N}$ such that $\mathcal{S}^{(s)}$ is an order α digital $(t(s), s)$ -sequence over \mathbb{F}_b for any $s \in \mathbb{N}$, then we call \mathcal{S} an infinite-dimensional order α digital sequence over \mathbb{F}_b .

Remark 1. It is clear from Definition 3 that any digital net can be regarded as an order α digital (t, m, s) -net with $t = \alpha m$ since then there is no linear independence condition imposed on the rows of the generating matrices. Once a digital sequence is considered, however, it follows from Definition 4 that the value of $t(s)$ is supposed to be independent of m . Since we are interested in the convergence behavior of the worst-case error as a function of m (instead of a fixed m), we allow the situation $\alpha m < t(s)$ and think of $P_{b^1}^{(s)}, P_{b^2}^{(s)}, \dots$ as a sequence of order α digital $(t(s), m, s)$ -nets with a fixed $t(s)$.

2.1.2 Explicit construction

Explicit constructions of infinite-dimensional order 1 digital sequences over \mathbb{F}_b have been given by Sobol' [34], Niederreiter [27], Tezuka [35], Niederreiter and Xing [29] and others. For instance, the Niederreiter sequence introduced in [27], which was generalized thereafter in [35], is constructed as follows. Let $p_1, p_2, \dots \in \mathbb{F}_b[x]$ be distinct monic irreducible polynomials over \mathbb{F}_b . For each $j \in \mathbb{N}$, let $e_j = \deg(p_j)$ and consider the following Laurent series expansion

$$\frac{x^{e_j - z - 1}}{p_j(x)^i} = \sum_{l=1}^{\infty} a^{(j)}(i, z, l) x^{-l} \in \mathbb{F}_b((x^{-1})),$$

for integers $i \geq 1$ and $0 \leq z < e_j$. Then define the matrix $C_j = (c_{k,l}^{(j)})_{k,l \in \mathbb{N}}$ by

$$c_{k,l}^{(j)} = a^{(j)}\left(\left\lfloor \frac{k-1}{e_j} \right\rfloor + 1, (k-1) \bmod e_j, l\right).$$

Note that we have $c_{k,l}^{(j)} = 0$ whenever $k > l$. It is known that these matrices C_j generate an infinite-dimensional order 1 digital sequence over \mathbb{F}_b whose function $t: \mathbb{N} \rightarrow \mathbb{N}$ is given by

$$t(s) = \sum_{j=1}^s (e_j - 1),$$

for any $s \in \mathbb{N}$. We refer to [11, Section 8] for more information on these special constructions of infinite-dimensional order 1 digital sequences. In what follows, we introduce the digit interlacing composition due to Dick [5, 6], which enables to explicitly construct infinite-dimensional order α digital sequences over \mathbb{F}_b for a given $\alpha \in \mathbb{N}$ by using infinite-dimensional order 1 digital sequences over \mathbb{F}_b .

Definition 5. For $\alpha \in \mathbb{N}$, let $\mathbf{x} = (x_1, \dots, x_\alpha) \in [0, 1)^\alpha$. For $1 \leq j \leq \alpha$, we denote the b -adic expansion of x_j by $x_j = \xi_{1,j}/b + \xi_{2,j}/b^2 + \dots$ with $\xi_{i,j} \in \mathbb{F}_b$, which is understood to be unique in the sense that infinitely many of the $\xi_{i,j}$'s are different from $b-1$. Then we define the map $\mathcal{D}_\alpha: [0, 1)^\alpha \rightarrow [0, 1)$ by

$$\mathcal{D}_\alpha(x_1, \dots, x_\alpha) := \sum_{i=1}^{\infty} \sum_{j=1}^{\alpha} \frac{\xi_{i,j}}{b^{\alpha(i-1)+j}}.$$

In case of an infinite-dimensional point $\mathbf{x} = (x_1, x_2, \dots) \in [0, 1]^\mathbb{N}$, we apply \mathcal{D}_α to every non-overlapping consecutive α components of \mathbf{x} , i.e.,

$$\mathcal{D}_\alpha(x_1, x_2, \dots) := (\mathcal{D}_\alpha(x_1, \dots, x_\alpha), \mathcal{D}_\alpha(x_{\alpha+1}, \dots, x_{2\alpha}), \dots) \in [0, 1]^\mathbb{N}.$$

By using infinite-dimensional order 1 digital sequences and \mathcal{D}_α , we can construct infinite-dimensional order α digital sequences as follows [6, Theorems 4.11 and 4.12].

Lemma 1. *Let $\mathcal{S} = \{\mathbf{x}_n : n \geq 0\} \subset [0, 1]^\mathbb{N}$ be an infinite-dimensional order 1 digital sequence over \mathbb{F}_b with $t = t_1 : \mathbb{N} \rightarrow \mathbb{N}$. For $\alpha \in \mathbb{N}$, define*

$$\mathcal{D}_\alpha(\mathcal{S}) := \{\mathcal{D}_\alpha(\mathbf{x}_n) : n \geq 0\} \subset [0, 1]^\mathbb{N}.$$

Then $\mathcal{D}_\alpha(\mathcal{S})$ is an infinite-dimensional order α digital sequence over \mathbb{F}_b with $t = t_\alpha : \mathbb{N} \rightarrow \mathbb{N}$, where t_α is given by

$$t_\alpha(s) = \alpha t_1(\alpha s) + \frac{s\alpha(\alpha - 1)}{2},$$

for any $s \in \mathbb{N}$.

Remark 2. *Let \mathcal{S} be an infinite-dimensional digital sequence with generating matrices $C_j \in \mathbb{F}_b^{\mathbb{N} \times \mathbb{N}}$. For $i, j \in \mathbb{N}$, let $\mathbf{c}_{i,j}$ denote the i -th row of C_j . For each $j \in \mathbb{N}$, define a matrix $D_j \in \mathbb{F}_b^{\mathbb{N} \times \mathbb{N}}$, whose i -th row is denoted by $\mathbf{d}_{i,j}$, as*

$$\mathbf{d}_{\alpha(h-1)+i,j} = \mathbf{c}_{h,\alpha(j-1)+i},$$

for $h \in \mathbb{N}$ and $1 \leq i \leq \alpha$. Then the infinite-dimensional digital sequence over \mathbb{F}_b with generating matrices $D_j \in \mathbb{F}_b^{\mathbb{N} \times \mathbb{N}}$ is nothing but $\mathcal{D}_\alpha(\mathcal{S})$ defined in Lemma 1.

2.1.3 Dual net and Dick metric functions

The concept of dual net is of crucial importance in analyzing the worst-case error of a digital net.

Definition 6. *For $m, n, s \in \mathbb{N}$, let $P_{b^m}^{(s)}$ be a digital net with generating matrices $C_1, \dots, C_s \in \mathbb{F}_b^{n \times m}$. Then the dual net of $P_{b^m}^{(s)}$, denoted by $P_{b^m}^{(s)\perp}$, is defined as*

$$P_{b^m}^{(s)\perp} := \left\{ \mathbf{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s : C_1^\top \vec{k}_1 \oplus \dots \oplus C_s^\top \vec{k}_s = \mathbf{0} \in \mathbb{F}_b^m \right\},$$

where we set $\vec{k} = (\kappa_1, \dots, \kappa_n)$ for $k \in \mathbb{N}_0$ whose b -adic expansion is denoted by $k = \sum_{i=1}^\infty \kappa_i b^{i-1}$, where all except a finite number of κ_i are 0.

Further we recall the definition of the Dick metric function μ_α for $\alpha \in \mathbb{N}$ [6]. Note that the special case where $\alpha = 1$ was originally introduced in [26] and [32], and is called the NRT metric function.

Definition 7. *Let $\alpha \in \mathbb{N}$. For $k \in \mathbb{N}$, we denote the b -adic expansion of k by $k = \kappa_1 b^{c_1-1} + \kappa_2 b^{c_2-1} + \dots + \kappa_v b^{c_v-1}$ such that $\kappa_1, \dots, \kappa_v \in \{1, \dots, b-1\}$ and $c_1 > c_2 > \dots > c_v > 0$. Then we define*

$$\mu_\alpha(k) := \sum_{i=1}^{\min(\alpha, v)} c_i,$$

and $\mu_\alpha(0) := 0$. For $\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s$, we define

$$\mu_\alpha(\mathbf{k}) := \sum_{j=1}^s \mu_\alpha(k_j).$$

Then the minimum Dick metric of a digital net is defined as follows.

Definition 8. For $m, n, s \in \mathbb{N}$, let $P_{b^m}^{(s)}$ be a digital net over \mathbb{F}_b and $P_{b^m}^{(s)\perp}$ its dual net. For $\alpha \in \mathbb{N}$, the minimum Dick metric of $P_{b^m}^{(s)}$ is defined by

$$\rho_\alpha(P_{b^m}^{(s)}) := \min_{\mathbf{k} \in P_{b^m}^{(s)\perp} \setminus \{\mathbf{0}\}} \mu_\alpha(\mathbf{k}).$$

2.1.4 Some properties of higher order digital nets and sequences

The following property of an order α digital (t, m, s) -net directly follows from the linear independence of the rows of generating matrices, see [11, Chapter 15].

Lemma 2. For any order α digital (t, m, s) -net $P_{b^m}^{(s)}$ over \mathbb{F}_b , we have

$$\alpha m - t < \rho_\alpha(P_{b^m}^{(s)}) \leq \alpha m.$$

Moreover, the following lemma is known under the name of *propagation rule*, which states that any order α digital net is also an order α' digital net as long as $1 \leq \alpha' < \alpha$. The lemma again directly follows from the linear independence of the rows of the generating matrices. We refer to [5, Theorem 3.3] and [6, Theorem 4.10] for the proof.

Lemma 3. For $\alpha \in \mathbb{N}$, let $P_{b^m}^{(s)}$ be an order α digital (t, m, s) -net over \mathbb{F}_b with some integer $0 \leq t \leq \alpha m$. Then, for any $\alpha' \in \mathbb{N}$ with $1 \leq \alpha' < \alpha$, $P_{b^m}^{(s)}$ is also an order α' digital $(t_{\alpha'}, m, s)$ -net over \mathbb{F}_b with $t_{\alpha'} = \lceil t\alpha'/\alpha \rceil$.

2.2 Sobolev spaces

Here we introduce the function space which we consider in this paper. First let us consider the one-dimensional case. For $\alpha \in \mathbb{N}$, the Sobolev space with smoothness α is given by

$$H_\alpha := \left\{ f: [0, 1) \rightarrow \mathbb{R} \mid \begin{array}{l} f^{(r)}: \text{absolutely continuous for } r = 0, \dots, \alpha - 1, \\ f^{(\alpha)} \in L^2([0, 1)) \end{array} \right\},$$

where $f^{(r)}$ denotes the r -th derivative of f , with the inner product

$$\langle f, g \rangle_\alpha = \sum_{r=0}^{\alpha-1} \int_0^1 f^{(r)}(x) \, dx \int_0^1 g^{(r)}(x) \, dx + \int_0^1 f^{(\alpha)}(x) g^{(\alpha)}(x) \, dx,$$

for $f, g \in H_\alpha$. The space H_α is also a reproducing kernel Hilbert space with the reproducing kernel

$$K_\alpha(x, y) = \sum_{r=0}^{\alpha} \frac{B_r(x) B_r(y)}{(r!)^2} + (-1)^{\alpha+1} \frac{B_{2\alpha}(|x-y|)}{(2\alpha)!},$$

for $x, y \in [0, 1)$, where B_r denotes the Bernoulli polynomial of degree r .

For the s -dimensional case, we consider the s -fold tensor product space of the one-dimensional space introduced above. Thus the Sobolev space $H_{\alpha,s}$ which we deal with is simply given by $H_{\alpha,s} = \bigotimes_{j=1}^s H_{\alpha}$. Then it follows from [1, Section 8] that the reproducing kernel of the space $H_{\alpha,s}$ is the product of the reproducing kernels for the one-dimensional space H_{α} . Therefore, $H_{\alpha,s}$ is the reproducing kernel Hilbert space with the inner product

$$\begin{aligned} \langle f, g \rangle_{\alpha,s} = & \sum_{u \subseteq \{1, \dots, s\}} \sum_{\mathbf{r}_u \in \{0, \dots, \alpha-1\}^{|u|}} \int_{[0,1]^{s-|u|}} \\ & \left(\int_{[0,1]^{|u|}} f^{(\mathbf{r}_u, \boldsymbol{\alpha})}(\mathbf{x}) \, d\mathbf{x}_u \right) \left(\int_{[0,1]^{|u|}} g^{(\mathbf{r}_u, \boldsymbol{\alpha})}(\mathbf{x}) \, d\mathbf{x}_u \right) d\mathbf{x}_{\{1, \dots, s\} \setminus u}, \end{aligned}$$

for $f, g \in H_{\alpha,s}$, and the reproducing kernel

$$K_{\alpha,s}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s K_{\alpha}(x_j, y_j),$$

for $\mathbf{x} = (x_1, \dots, x_s), \mathbf{y} = (y_1, \dots, y_s) \in [0, 1]^s$. In the above, we use the following notation: For $u \subseteq \{1, \dots, s\}$ and $\mathbf{x} \in [0, 1]^s$, we write $\mathbf{x}_u = (x_j)_{j \in u}$. Moreover, for $\mathbf{r}_u = (r_j)_{j \in u} \in \{0, \dots, \alpha-1\}^{|u|}$, $(\mathbf{r}_u, \boldsymbol{\alpha})$ denotes the s -dimensional vector whose j -th component equals r_j if $j \in u$, and α otherwise. Note that an integral and sum over the empty set is defined to be the identity operator.

2.3 Walsh functions

Here we recall the definition of Walsh functions. First let us define the one-dimensional Walsh functions.

Definition 9. For $b \in \mathbb{N}$, $b \geq 2$, let $\omega_b := \exp(2\pi\sqrt{-1}/b)$. For $k \in \mathbb{N}_0$, we denote the b -adic expansion of k by $k = \sum_{i=1}^{\infty} \kappa_i b^{i-1}$, where all except a finite number of κ_i are 0. The k -th b -adic Walsh function ${}_b\text{wal}_k: [0, 1) \rightarrow \{1, \omega_b, \dots, \omega_b^{b-1}\}$ is defined by

$${}_b\text{wal}_k(x) := \omega_b^{\kappa_1 \xi_1 + \kappa_2 \xi_2 + \dots},$$

where we denote the unique b -adic expansion of $x \in [0, 1)$ by $x = \sum_{i=1}^{\infty} \xi_i b^{-i}$ with $\xi_i \in \mathbb{F}_b$.

The above definition can be extended to the multi-variate case as follows.

Definition 10. For $b \in \mathbb{N}$, $b \geq 2$ and $\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s$, the \mathbf{k} -th b -adic Walsh function ${}_b\text{wal}_{\mathbf{k}}: [0, 1)^s \rightarrow \{1, \omega_b, \dots, \omega_b^{b-1}\}$ is defined by

$${}_b\text{wal}_{\mathbf{k}}(\mathbf{x}) := \prod_{j=1}^s {}_b\text{wal}_{k_j}(x_j).$$

Since we shall always use Walsh functions in a fixed prime base b , we omit the subscript and simply write wal_k or $\text{wal}_{\mathbf{k}}$.

As shown in [11, Theorem A.11], the Walsh system $\{\text{wal}_{\mathbf{k}} : \mathbf{k} \in \mathbb{N}_0^s\}$ is a complete orthonormal system in $L^2([0, 1]^s)$ for any $s \in \mathbb{N}$. Thus, we can define the Walsh series of $f \in L^2([0, 1]^s)$ by

$$\sum_{\mathbf{k} \in \mathbb{N}_0^s} \hat{f}(\mathbf{k}) \text{wal}_{\mathbf{k}}(\mathbf{x}),$$

where $\hat{f}(\mathbf{k})$ denotes the \mathbf{k} -th Walsh coefficient of f defined by

$$\hat{f}(\mathbf{k}) := \int_{[0, 1]^s} f(\mathbf{x}) \overline{\text{wal}_{\mathbf{k}}(\mathbf{x})} d\mathbf{x}.$$

Regarding the reproducing kernel $K_{\alpha, s}$ given in the last subsection, we have

$$\begin{aligned} \hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l}) &:= \int_{[0, 1]^{2s}} K_{\alpha, s}(\mathbf{x}, \mathbf{y}) \overline{\text{wal}_{\mathbf{k}}(\mathbf{x})} \text{wal}_{\mathbf{l}}(\mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \int_{[0, 1]^{2s}} \prod_{j=1}^s K_{\alpha}(x_j, y_j) \overline{\text{wal}_{k_j}(x_j)} \text{wal}_{l_j}(y_j) d\mathbf{x} d\mathbf{y} \\ &= \prod_{j=1}^s \int_{[0, 1]^2} K_{\alpha}(x_j, y_j) \overline{\text{wal}_{k_j}(x_j)} \text{wal}_{l_j}(y_j) dx_j dy_j = \prod_{j=1}^s \hat{K}_{\alpha}(k_j, l_j), \end{aligned} \tag{1}$$

for any $\mathbf{k} = (k_1, \dots, k_s), \mathbf{l} = (l_1, \dots, l_s) \in \mathbb{N}_0^s$. As shown in [2, Proposition 20], there exists a positive constant $D_{\alpha, b}$ such that

$$\left| \hat{K}_{\alpha}(k, l) \right| \leq D_{\alpha, b} b^{-\mu_{\alpha}(k) - \mu_{\alpha}(l)} \tag{2}$$

for any $k, l \in \mathbb{N}_0$. In order to ensure the pointwise absolute convergence of the Walsh series of $K_{\alpha, s}$, it suffices to prove that $K_{\alpha, s}$ is continuous, which is in fact trivial, and that

$$\sum_{\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s} \left| \hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l}) \right| < \infty. \tag{3}$$

To address the proof of our main result first, we postpone the proof of (3) to Section 4, where the result in Subsection 3.1 shall be used when $\alpha = 1$. Nevertheless, for any $\alpha \in \mathbb{N}$, the following pointwise equality holds

$$\begin{aligned} K_{\alpha, s}(\mathbf{x}, \mathbf{y}) &= \sum_{\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s} \hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l}) \text{wal}_{\mathbf{k}}(\mathbf{x}) \overline{\text{wal}_{\mathbf{l}}(\mathbf{y})} \\ &= \sum_{\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s} \left(\prod_{j=1}^s \hat{K}_{\alpha}(k_j, l_j) \right) \text{wal}_{\mathbf{k}}(\mathbf{x}) \overline{\text{wal}_{\mathbf{l}}(\mathbf{y})}. \end{aligned}$$

3 Proof of the main result

We repeat the main result of this paper for the reader's convenience:

Theorem 1. *Let b be a prime, and $s, \alpha \in \mathbb{N}$. Let \mathcal{S} be an infinite-dimensional order $2\alpha + 1$ digital sequence over the finite field \mathbb{F}_b , and let $P_N^{(s)}$ be the set of the first N points of the projection of \mathcal{S} onto the first s coordinates for $N \in \mathbb{N}$. Then for any $m \in \mathbb{N}$, we have*

$$e^{\text{wor}}(H_{\alpha,s}; P_{b^m}^{(s)}) \leq C_{\alpha,b,s} \frac{m^{(s-1)/2}}{b^{\alpha m}},$$

where $C_{\alpha,b,s}$ is positive and independent of m .

Again we recall that the infinite-dimensional order $2\alpha + 1$ digital sequence \mathcal{S} can be constructed by applying the digit interlacing composition $\mathcal{D}_{2\alpha+1}$ to the infinite-dimensional order 1 digital sequence \mathcal{S}_1 . Thus the point set $P_{b^m}^{(s)}$ is obtained as

$$\mathcal{S}_1 \rightarrow \mathcal{D}_{2\alpha+1}(\mathcal{S}_1) = \mathcal{S} \rightarrow P_{b^m}^{(s)}.$$

Throughout this section we write P instead of $P_{b^m}^{(s)}$ for ease of notation, and denote the dual net of P by P^\perp . As stated in Remark 1, P is an order $2\alpha + 1$ digital (t, m, s) -net with t independent of m . Moreover, for $f, g: \mathbb{N} \rightarrow \mathbb{R}$, we write $f(m) \ll_{a,b} g(m)$ if there exists a positive constant C which depends on some parameters (a and b in this case) such that $f(m) \leq Cg(m)$ for all m . We first give the proof of Theorem 1 by using Lemma 8, which is shown later.

Proof of Theorem 1. In order to prove Theorem 1, it suffices from Definition 4 and Remark 1 to prove that the inequality

$$e^{\text{wor}}(H_{\alpha,s}; P_{b^m}^{(s)}) \ll_{\alpha,b,s,t} \frac{m^{(s-1)/2}}{b^{\alpha m}} \quad (4)$$

holds for any $m \in \mathbb{N}$.

Using [2, Theorem 15], (1) and (2), we have

$$\begin{aligned} (e^{\text{wor}}(H_{\alpha,s}; P))^2 &= \sum_{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \\ &\leq \sum_{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \left| \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \right| = \sum_{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}} \prod_{j=1}^s \left| \hat{K}_\alpha(k_j, l_j) \right| \\ &\ll_{\alpha,b,s} \sum_{\substack{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \neq 0}} \prod_{j=1}^s b^{-\mu_\alpha(k_j) - \mu_\alpha(l_j)} \\ &= \sum_{\substack{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \neq 0}} b^{-\mu_\alpha(\mathbf{k}) - \mu_\alpha(\mathbf{l})}. \end{aligned}$$

Let us recall the interpolation inequality for Dick metric functions given in [15, Lemma 3.1], which states that for $\alpha, \beta \in \mathbb{N}$ with $1 < \alpha \leq \beta$ we have

$$\mu_\alpha(\mathbf{k}) \geq \frac{\alpha-1}{\beta-1} \mu_\beta(\mathbf{k}) + \frac{\beta-\alpha}{\beta-1} \mu_1(\mathbf{k}),$$

for any $\mathbf{k} \in \mathbb{N}_0^s$. Since we only consider the case $\beta = 2\alpha + 1$ in the following, we simply write

$$\mu_\alpha(\mathbf{k}) \geq A\mu_{2\alpha+1}(\mathbf{k}) + B\mu_1(\mathbf{k}), \quad (5)$$

where $A = (\alpha - 1)/(2\alpha)$ and $B = (\alpha + 1)/(2\alpha)$. It follows from the inequality (5) and Definition 8 that

$$\begin{aligned} (e^{\text{wor}}(H_{\alpha,s}; P))^2 &\ll_{\alpha,b,s} \sum_{\substack{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \neq 0}} b^{-\mu_\alpha(\mathbf{k}) - \mu_\alpha(\mathbf{l})} \\ &\leq \sum_{\substack{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \neq 0}} b^{-A(\mu_{2\alpha+1}(\mathbf{k}) + \mu_{2\alpha+1}(\mathbf{l})) - B(\mu_1(\mathbf{k}) + \mu_1(\mathbf{l}))} \\ &\leq b^{-2A\rho_{2\alpha+1}(P)} \sum_{\substack{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \neq 0}} b^{-B(\mu_1(\mathbf{k}) + \mu_1(\mathbf{l}))} \\ &= b^{-2A\rho_{2\alpha+1}(P)} \sum_{z=2\rho_1(P)}^{\infty} \sum_{\substack{\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} \\ \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \neq 0, \mu_1(\mathbf{k}) + \mu_1(\mathbf{l}) = z}} b^{-Bz}, \\ &= b^{-2A\rho_{2\alpha+1}(P)} \sum_{z=2\rho_1(P)}^{\infty} \frac{|J_\alpha(z)|}{b^{Bz}}, \end{aligned} \quad (6)$$

where we define

$$J_\alpha(z) := \{(\mathbf{k}, \mathbf{l}) \in (P^\perp \setminus \{\mathbf{0}\})^2 : \hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l}) \neq 0, \mu_1(\mathbf{k}) + \mu_1(\mathbf{l}) = z\}. \quad (7)$$

Plugging the bound on $|J_\alpha(z)|$ which will be given in Lemma 8, i.e.,

$$|J_\alpha(z)| \ll_{\alpha,b,s,t} (z - 2\rho_1(P))^{2s\alpha+1} z^{s-1} b^{(z-2\rho_1(P))/2},$$

into (6) and using the change of variables $z = \kappa + 2\rho_1(P)$, we obtain

$$\begin{aligned} (e^{\text{wor}}(H_{\alpha,s}; P))^2 &\ll_{\alpha,b,s,t} b^{-2A\rho_{2\alpha+1}(P)} \sum_{\kappa=0}^{\infty} \frac{\kappa^{2\alpha s+1} (\kappa + 2\rho_1(P))^{s-1} b^{\kappa/2}}{b^{B(\kappa+2\rho_1(P))}} \\ &\ll_s \frac{\rho_1(P)^{s-1}}{b^{2A\rho_{2\alpha+1}(P)+2B\rho_1(P)}} \sum_{\kappa=0}^{\infty} \frac{\kappa^{s(2\alpha+1)}}{b^{\kappa/(2\alpha)}} \\ &\ll_{\alpha,b,s} \frac{m^{s-1}}{b^{2A\rho_{2\alpha+1}(P)+2B\rho_1(P)}}, \end{aligned}$$

where the last inequality stems from the facts that the sum over κ is finite and depends only on s, α, b and that $\rho_1(P)$ cannot be larger than m . Finally, since it follows from Lemmas 2 and 3 that

$$\begin{aligned} 2A\rho_{2\alpha+1}(P) + 2B\rho_1(P) &\geq \frac{\alpha-1}{\alpha} ((2\alpha+1)m - t) + \frac{\alpha+1}{\alpha} (m - \lceil t/(2\alpha+1) \rceil) \\ &= 2\alpha m - \frac{\alpha-1}{\alpha} t - \frac{\alpha+1}{\alpha} \lceil t/(2\alpha+1) \rceil, \end{aligned}$$

we have

$$(e^{\text{wor}}(H_{\alpha,s}; P))^2 \ll_{\alpha,b,s,t} \frac{m^{s-1}}{b^{2\alpha m}}.$$

Thus the result follows by taking the square root. \square

3.1 Sparsity of the Walsh coefficients

Here we show the sparsity of the Walsh coefficients $\hat{K}_{\alpha,s}$ in a stronger form than that given by the authors in [16, Section 4.1]. Let us consider the one-dimensional case first. As can be seen from [2, Section 3.1], the Walsh coefficient of the univariate reproducing kernel K_α is given by

$$\hat{K}_\alpha(k, l) = \sum_{r=0}^{\alpha} \hat{b}_r(k) \overline{\hat{b}_r(l)} + (-1)^{\alpha+1} \hat{b}_{2\alpha, \text{per}}(k, l), \quad (8)$$

for $k, l \in \mathbb{N}_0$, where we write

$$\begin{aligned} \hat{b}_r(k) &:= \int_0^1 \frac{B_r(x)}{r!} \overline{\text{wal}_k(x)} dx, \\ \hat{b}_{r, \text{per}}(k, l) &:= \int_0^1 \int_0^1 \frac{\tilde{B}_r(|x-y|)}{r!} \overline{\text{wal}_k(x)} \text{wal}_l(y) dx dy, \end{aligned}$$

where $\tilde{B}_r : \mathbb{R} \rightarrow \mathbb{R}$ is defined by extending B_r periodically to \mathbb{R} . Note that we have $B_r(|x-y|) = \tilde{B}_r(x-y)$ for even r and $B_r(|x-y|) = (-1)^{1_{x < y}} \tilde{B}_r(x-y)$ for odd r for any $x, y \in [0, 1)$, where $1_{x < y}$ equals 1 if $x < y$ and 0 otherwise.

In order to prove that the Walsh coefficients \hat{K}_α are sparse, i.e., $\hat{K}_\alpha(k, l) = 0$ for many choices of $k, l \in \mathbb{N}_0$, we introduce the notion of type (p, q) .

Definition 11. For $k, l \in \mathbb{N}_0$, we denote the b -adic expansions of k and l by

$$k = \sum_{i=1}^v \kappa_i b^{c_i-1} \quad \text{and} \quad l = \sum_{i=1}^w \lambda_i b^{d_i-1},$$

respectively, where $\kappa_1, \dots, \kappa_v, \lambda_1, \dots, \lambda_w \in \{1, \dots, b-1\}$, $c_1 > c_2 > \dots > c_v > 0$ and $d_1 > d_2 > \dots > d_w > 0$. For $k = 0$ ($l = 0$, resp.), we assume that $v = 0$ and $\kappa_0 b^{c_0-1} = 0$ ($w = 0$ and $\lambda_0 b^{d_0-1} = 0$, resp.). For $p, q \in \mathbb{N}_0$, we write

$$k^{(p)} = \sum_{i=p+1}^v \kappa_i b^{c_i-1} \quad \text{and} \quad l^{(q)} = \sum_{i=q+1}^w \lambda_i b^{d_i-1},$$

where the empty sum equals 0. Then we say that (k, l) is of type (p, q) if $k^{(p)} = l^{(q)}$ and $\kappa_p b^{c_p-1} \neq \lambda_q b^{d_q-1}$, where we set $\kappa_0 b^{c_0-1} = \lambda_0 b^{d_0-1} = 0$, except the case $k = l$ where we say that (k, l) is of type $(0, 0)$.

Remark 3. If (k, l) is of type (p, q) , then it follows from the above definition that $v - p = w - q$, and that $\kappa_{v-i} = \lambda_{w-i}$ and $c_{v-i} = d_{w-i}$ for all $0 \leq i < v - p$. For any $k, l \in \mathbb{N}_0$, there is a unique set of nonnegative integers p, q with $p \leq v$ and $q \leq w$ such that (k, l) is of type (p, q) .

Now the sparsity of the Walsh coefficients can be formulated as follows.

Proposition 1. Let $k, l \in \mathbb{N}_0$ such that (k, l) is of type (p, q) with $p + q > 2\alpha$. Then we have $\hat{K}_\alpha(k, l) = 0$.

In order to prove Proposition 1, it suffices to show that every term on the right-hand side of (8) is 0 whenever (k, l) is of type (p, q) such that $p + q > 2\alpha$, which shall be proven in Lemmas 4 and 5 below.

Lemma 4. *Let $k, l \in \mathbb{N}_0$ such that (k, l) is of type (p, q) with $p + q > 2\alpha$. Then we have $\hat{b}_r(k)\hat{b}_r(l) = 0$ for all $0 \leq r \leq \alpha$.*

Proof. Since $p + q > 2\alpha$, we must have either $p > \alpha$ or $q > \alpha$, which implies that either $v > \alpha$ or $w > \alpha$ holds, i.e., the b -adic expansion of either k or l contains more than α non-zero terms. Then it is known from [6, Lemma 3.7] that either the k -th or l -th Walsh coefficients of polynomials of degree less than or equal to α are all 0. This means either $\hat{b}_r(k) = 0$ or $\hat{b}_r(l) = 0$ for all $0 \leq r \leq \alpha$, which completes the proof. \square

Lemma 5. *For $r \in \mathbb{N}$, $r \geq 2$, let $k, l \in \mathbb{N}_0$ such that (k, l) is of type (p, q) with $p + q > r$. Then we have $\hat{b}_{r,\text{per}}(k, l) = 0$.*

Proof. We prove this lemma by induction on $r \geq 2$.

Let us consider the case $r = 2$ first. If (k, l) is of type (p, q) with $p + q > 2$, it never follows that $k = l$, $k^{(1)} = l^{(1)}$ with $k \neq l$, $k^{(1)} = l$, $k = l^{(1)}$, $k^{(2)} = l$, or $k = l^{(2)}$, since in such cases (k, l) becomes of type (p, q) with $p + q = 0, 2, 1, 1, 2$ and 2 , respectively. Then the result immediately follows from [7, Lemma 10].

Now assume that the result holds true for $r - 1$. That is, we assume that

$$\hat{b}_{r-1,\text{per}}(k, l) = 0 \quad \text{for } k, l \in \mathbb{N}_0 \text{ of type } (p, q) \text{ with } p + q > r - 1. \quad (9)$$

If either $k = 0$ or $l = 0$ holds, the result $\hat{b}_{r,\text{per}}(k, l) = 0$ immediately follows from [7, Lemma 11]. Thus we focus on the case $k, l > 0$ in the following. Moreover, since $\hat{b}_{r,\text{per}}(l, k) = \overline{\hat{b}_{r,\text{per}}(k, l)}$, we can restrict ourselves to the case $p > r/2$. As in [11, Equation 14.18], for any $k, l \in \mathbb{N}$ and $r > 2$ we have the identity

$$\begin{aligned} \hat{b}_{r,\text{per}}(k, l) = & -\frac{1}{b^{c_1}} \left(\frac{1}{1 - \omega_b^{-\kappa_1}} \hat{b}_{r-1,\text{per}}(k^{(1)}, l) + \left(\frac{1}{2} + \frac{1}{\omega_b^{-\kappa_1} - 1} \right) \hat{b}_{r-1,\text{per}}(k, l) \right. \\ & \left. + \sum_{a=1}^{\infty} \sum_{\theta=1}^{b-1} \frac{1}{b^a(\omega_b^{\theta} - 1)} \hat{b}_{r-1,\text{per}}(\theta b^{a+c_1-1} + k, l) \right). \end{aligned}$$

Thus, in order to prove $\hat{b}_{r,\text{per}}(k, l) = 0$ for $k, l \in \mathbb{N}$ of type (p, q) with $p + q > r$, it suffices to prove that (i) $\hat{b}_{r-1,\text{per}}(k^{(1)}, l) = 0$, (ii) $\hat{b}_{r-1,\text{per}}(k, l) = 0$ and (iii) $\hat{b}_{r-1,\text{per}}(\theta b^{a+c_1-1} + k, l) = 0$ for any $a \in \mathbb{N}$ and $1 \leq \theta \leq b-1$ whenever $p + q > r$ and $p > r/2$.

First, from the induction assumption (9), it is trivial that $\hat{b}_{r-1,\text{per}}(k, l) = 0$ also for $k, l \in \mathbb{N}$ of type (p, q) with $p + q > r$. The remaining two items (i) and (iii) can be proven in the following way.

As $k^{(1)}$ can be written as

$$k^{(1)} = \sum_{i=2}^p \kappa_i b^{c_i-1} + \sum_{i=q+1}^w \lambda_i b^{d_i-1},$$

where $\kappa_p b^{c_p-1} \neq \lambda_q b^{d_q-1}$, $(k^{(1)}, l)$ is of type $(p-1, q)$, where $p-1+q > r-1$. Thus, again from the induction assumption (9), it follows that $\hat{b}_{r-1,\text{per}}(k^{(1)}, l) = 0$, which completes the proof of the item (i).

Similarly, as $\theta b^{a+c_1-1} + k$ can be written as

$$\theta b^{a+c_1-1} + k = \theta b^{a+c_1-1} + \sum_{i=1}^p \kappa_i b^{c_i-1} + \sum_{i=q+1}^w \lambda_i b^{d_i-1},$$

where $\kappa_p b^{c_p-1} \neq \lambda_q b^{d_q-1}$, for any $a \in \mathbb{N}$ and $1 \leq \theta \leq b-1$, $(\theta b^{a+c_1-1} + k, l)$ is of type $(p+1, q)$, where $p+1+q > r+1$. Again from the induction assumption (9), it follows that $\hat{b}_{r-1, \text{per}}(\theta b^{c+a_1-1} + k, l) = 0$, which completes the proof of the item (iii). \square

Let us move on to the high-dimensional case. As a corollary of Proposition 1, the sparsity of the Walsh coefficients $\hat{K}_{\alpha, s}$ can be formulated as follows.

Corollary 1. *Let $\mathbf{k} = (k_1, \dots, k_s), \mathbf{l} = (l_1, \dots, l_s) \in \mathbb{N}_0$ such that (k_j, l_j) is of type (p_j, q_j) with $p_j + q_j > 2\alpha$ for at least one index $j \in \{1, \dots, s\}$. Then we have $\hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l}) = 0$.*

Proof. For an index $j \in \{1, \dots, s\}$ such that (k_j, l_j) is of type (p_j, q_j) with $p_j + q_j > 2\alpha$, it follows from Proposition 1 that $\hat{K}_{\alpha}(\mathbf{k}_j, \mathbf{l}_j) = 0$. The result then follows from (1). \square

3.2 A bound on $|J_{\alpha}(z)|$

Here we give a bound on the cardinality of $J_{\alpha}(z)$ defined in (7) by using the result on the sparsity of the Walsh coefficients $\hat{K}_{\alpha, s}$ and the property of higher order digital nets. For this purpose, we introduce the following two sets

$$\begin{aligned} R_{\alpha}(\mathbf{k}, z_2) &:= \{\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\} : \mu_1(\mathbf{l}) = z_2, \hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l}) \neq 0\}, \\ J_{\alpha}(z_1, z_2) &:= \{(\mathbf{k}, \mathbf{l}) \in (P^{\perp} \setminus \{\mathbf{0}\})^2 : \mu_1(\mathbf{k}) = z_1, \mu_1(\mathbf{l}) = z_2, \hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l}) \neq 0\}, \end{aligned}$$

for $z_1, z_2 \in \mathbb{N}$ and $\mathbf{k} \in \mathbb{N}_0^s$. In what follows, $\binom{i}{j}$ denotes the binomial coefficient, where we set $\binom{i}{j} = 0$ if $j > i$.

We first give a bound on $|R_{\alpha}(\mathbf{k}, z_2)|$. It is obvious from Definition 8 that $\mu_1(\mathbf{l}) \geq \rho_1(P)$ for $\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\}$. Thus $|R_{\alpha}(\mathbf{k}, z_2)| = 0$ for $z_2 < \rho_1(P)$. For $z_2 \geq \rho_1(P)$ the following holds true.

Lemma 6. *Let P be an order $2\alpha+1$ digital (t, m, s) -net over \mathbb{F}_b . For $z_2 \geq \rho_1(P)$ and $\mathbf{k} \in \mathbb{N}_0^s$, we have*

$$|R_{\alpha}(\mathbf{k}, z_2)| \leq_{\alpha, b, s} \prod_{i=1}^{2\alpha} \binom{(i+1)z_2 - \rho_{i+1}(P) + s}{s}.$$

Proof. For $\mathbf{p} = (p_1, \dots, p_s), \mathbf{q} = (q_1, \dots, q_s) \in \mathbb{N}_0^s$, define the set

$$R_{\mathbf{p}, \mathbf{q}}(\mathbf{k}, z_2) := \{\mathbf{l} \in P^{\perp} \setminus \{\mathbf{0}\} : \mu_1(\mathbf{l}) = z_2, (k_j, l_j) \text{ is of type } (p_j, q_j)\}.$$

It follows from Corollary 1 that in order to have $\hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l}) \neq 0$, (k_j, l_j) must be of type (p_j, q_j) with $p_j + q_j \leq 2\alpha$ for all $j = 1, \dots, s$. Thus we have

$$R_{\alpha}(\mathbf{k}, z_2) \subset \bigcup_{\substack{\mathbf{p}, \mathbf{q} \in \mathbb{N}_0^s \\ p_j + q_j \leq 2\alpha, \forall j}} R_{\mathbf{p}, \mathbf{q}}(\mathbf{k}, z_2). \quad (10)$$

Note that each $R_{\mathbf{p}, \mathbf{q}}(\mathbf{k}, z_2)$ is a subset of P^{\perp} and that the above union is disjoint. Further define the set

$$S_{2\alpha}(z_2) := \left\{ (\delta_{i,j}, \zeta_{i,j})_{1 \leq i \leq 2\alpha, 1 \leq j \leq s} \in (\mathbb{N}_0 \times \{1, \dots, b-1\})^{2s\alpha} \right\}$$

$$: \sum_{j=1}^s \delta_{i,j} \leq (i+1)z_2 - \rho_{i+1}(P), \quad i = 1, \dots, 2\alpha \},$$

and the mapping $\phi_{2\alpha}$ from \mathbb{N}_0^s to $(\mathbb{N}_0 \times \{1, \dots, b-1\})^{2s\alpha}$ by

$$(l_j)_{1 \leq j \leq s} \mapsto (d_{i,j} - d_{i+1,j}, \lambda_{i,j})_{1 \leq i \leq 2\alpha, 1 \leq j \leq s},$$

where we denote the b -adic expansion of l_j by $l_j = \sum_{i=1}^{w_j} \lambda_{i,j} b^{d_{i,j}-1}$. For $i > w_j$, we set $(d_{i,j}, \lambda_{i,j}) = (0, 1)$.

First we show that the image of the restriction of $\phi_{2\alpha}$ to $R_{\mathbf{p},\mathbf{q}}(\mathbf{k}, z_2)$ is included in $S_{2\alpha}(z_2)$ for any $\mathbf{p}, \mathbf{q} \in \mathbb{N}_0^s$ with $p_j + q_j \leq 2\alpha$. From Definition 8 we have

$$\mu_i(\mathbf{l}) = \sum_{j=1}^s \sum_{l=1}^i d_{l,j} \geq \rho_i(P),$$

for any $i \in \mathbb{N}$ and $\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}$. Thus it follows that again for any $i \in \mathbb{N}$

$$\begin{aligned} \sum_{j=1}^s (d_{i,j} - d_{i+1,j}) &= \sum_{j=1}^s d_{i,j} - \left(\sum_{j=1}^s \sum_{l=1}^{i+1} d_{l,j} - \sum_{j=1}^s \sum_{l=1}^i d_{l,j} \right) \\ &\leq \sum_{j=1}^s d_{i,j} - \rho_{i+1}(P) + \sum_{j=1}^s \sum_{l=1}^i d_{l,j} \\ &\leq \sum_{j=1}^s d_{1,j} - \rho_{i+1}(P) + \sum_{j=1}^s i d_{1,j} \\ &= (i+1)\mu_1(\mathbf{l}) - \rho_{i+1}(P) = (i+1)z_2 - \rho_{i+1}(P). \end{aligned}$$

Thus it implies that $(d_{i,j} - d_{i+1,j}, \lambda_{i,j})_{1 \leq i \leq 2\alpha, 1 \leq j \leq s} \in S_{2\alpha}(z_2)$.

Next we show that the restricted map $\phi_{2\alpha}|_{R_{\mathbf{p},\mathbf{q}}(\mathbf{k}, z_2)}: R_{\mathbf{p},\mathbf{q}}(\mathbf{k}, z_2) \rightarrow S_{2\alpha}(z_2)$ is injective. We denote the b -adic expansion of k_j by $k_j = \sum_{i=1}^{v_j} \kappa_{i,j} b^{c_{i,j}-1}$ for $\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s$. Let $\mathbf{l} \in R_{\mathbf{p},\mathbf{q}}(\mathbf{k}, z_2)$ and $\phi_{2\alpha}(\mathbf{l}) = (\delta_{i,j}, \zeta_{i,j})_{1 \leq i \leq 2\alpha, 1 \leq j \leq s}$. It suffices to show that \mathbf{l} is determined by $(\delta_{i,j}, \zeta_{i,j})_{1 \leq i \leq 2\alpha, 1 \leq j \leq s}$, \mathbf{k} , \mathbf{p} and \mathbf{q} . Since (k_j, l_j) is of type (p_j, q_j) , we have

$$\sum_{i=q_j+1}^{w_j} \lambda_{i,j} b^{d_{i,j}-1} = \sum_{i=p_j+1}^{v_j} \kappa_{i,j} b^{c_{i,j}-1},$$

where the empty sum equals 0, and $d_{q_j+1,j} = c_{p_j+1,j}$. Note that if $q_j = w_j$, which is equivalent to $p_j = v_j$, then we set $d_{w_j+1,j} = c_{v_j+1,j} = 0$. Further, by the definition of $\phi_{2\alpha}$ and the fact $q_j \leq 2\alpha$, for all $1 \leq i \leq q_j$ we have

$$\lambda_{i,j} = \zeta_{i,j} \quad \text{and} \quad d_{i,j} = \sum_{h=i}^{q_j} \delta_{h,j} + d_{q_j+1,j} = \sum_{h=i}^{q_j} \delta_{h,j} + c_{p_j+1,j}.$$

Thus l_j can be determined by $(\delta_{i,j}, \zeta_{i,j})_{1 \leq i \leq 2\alpha}$, k_j , p_j and q_j as

$$l_j = \sum_{i=1}^{w_j} \lambda_{i,j} b^{d_{i,j}-1} = \sum_{i=1}^{q_j} \zeta_{i,j} b^{c_{p_j+1,j} + \sum_{h=i}^{q_j} \delta_{h,j}-1} + \sum_{i=p_j+1}^{v_j} \kappa_{i,j} b^{c_{i,j}-1},$$

where the empty sum equals 0. This implies the injectivity of $\phi_{2\alpha}|_{R_{\mathbf{p},\mathbf{q}}(\mathbf{k},z_2)}$.

By using (10) and the result that $\phi_{2\alpha}|_{R_{\mathbf{p},\mathbf{q}}(\mathbf{k},z_2)}: R_{\mathbf{p},\mathbf{q}}(\mathbf{k},z_2) \rightarrow S_{2\alpha}(z_2)$ is injective, we obtain

$$\begin{aligned} |R_\alpha(\mathbf{k}, z_2)| &\leq \sum_{\substack{\mathbf{p}, \mathbf{q} \in \mathbb{N}_0^s \\ p_j + q_j \leq 2\alpha, \forall j}} |R_{\mathbf{p},\mathbf{q}}(\mathbf{k}, z_2)| \leq \sum_{\substack{\mathbf{p}, \mathbf{q} \in \mathbb{N}_0^s \\ p_j + q_j \leq 2\alpha, \forall j}} |S_{2\alpha}(z_2)| \\ &= ((\alpha + 1)(2\alpha + 1))^s |S_{2\alpha}(z_2)| \ll_{\alpha, s} |S_{2\alpha}(z_2)|, \end{aligned}$$

where the number $(\alpha + 1)(2\alpha + 1)$ represents the total number of possible combinations of $p, q \in \mathbb{N}_0$ such that $p + q \leq 2\alpha$. Finally the cardinality of $S_{2\alpha}(z_2)$ is given by

$$\begin{aligned} |S_{2\alpha}(z_2)| &= (b-1)^{2\alpha s} \prod_{i=1}^{2\alpha} \left| \left\{ (\delta_{i,j})_{1 \leq j \leq s} \in \mathbb{N}_0^s : \sum_{j=1}^s \delta_{i,j} \leq (i+1)z_2 - \rho_{i+1}(P) \right\} \right| \\ &= (b-1)^{2\alpha s} \prod_{i=1}^{2\alpha} \sum_{r_i=0}^{(i+1)z_2 - \rho_{i+1}(P)} \left| \left\{ (\delta_{i,j})_{1 \leq j \leq s} \in \mathbb{N}_0^s : \sum_{j=1}^s \delta_{i,j} = r_i \right\} \right| \\ &= (b-1)^{2\alpha s} \prod_{i=1}^{2\alpha} \sum_{r_i=0}^{(i+1)z_2 - \rho_{i+1}(P)} \binom{r_i + s - 1}{s - 1} \\ &= (b-1)^{2\alpha s} \prod_{i=1}^{2\alpha} \binom{(i+1)z_2 - \rho_{i+1}(P) + s}{s}, \end{aligned}$$

which completes the proof. \square

We move on to a bound on $|J_\alpha(z_1, z_2)|$. It is again obvious from Definition 8 that $\mu_1(\mathbf{k}), \mu_1(\mathbf{l}) \geq \rho_1(P)$ for $\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}$. Thus $|J_\alpha(z_1, z_2)| = 0$ when either $z_1 < \rho_1(P)$ or $z_2 < \rho_1(P)$. For $z_1, z_2 \geq \rho_1(P)$ the following holds true.

Lemma 7. *Let P be an order $2\alpha + 1$ digital (t, m, s) -net over \mathbb{F}_b . For $z_1, z_2 \geq \rho_1(P)$, we have*

$$|J_\alpha(z_1, z_2)| \ll_{\alpha, b, s} b^{z_1 - \rho_1(P) + 1} \binom{z_1 + s - 1}{s - 1} \prod_{i=1}^{2\alpha} \binom{(i+1)z_2 - \rho_{i+1}(P) + s}{s}.$$

Proof. From the definitions of $R_\alpha(\mathbf{k}, z_2)$ and $J_\alpha(z_1, z_2)$, we have

$$\begin{aligned} |J_\alpha(z_1, z_2)| &\leq |\{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_1(\mathbf{k}) = z_1\}| \\ &\quad \times \max_{\substack{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} \\ \mu_1(\mathbf{k}) = z_1}} |\{\mathbf{l} \in P^\perp \setminus \{\mathbf{0}\} : \mu_1(\mathbf{l}) = z_2, \hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l}) \neq 0\}| \\ &= |I(z_1)| \max_{\substack{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} \\ \mu_1(\mathbf{k}) = z_1}} |R_\alpha(\mathbf{k}, z_2)|, \end{aligned}$$

where we write

$$I(z_1) := \{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_1(\mathbf{k}) = z_1\}.$$

By adapting [33, Lemma 2.2], we have

$$\begin{aligned}
|I(z_1)| &= |\{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_1(\mathbf{k}) = z_1\}| \\
&= \sum_{\substack{d_1, \dots, d_s \in \mathbb{N}_0 \\ d_1 + \dots + d_s = z_1}} |\{\mathbf{k} \in P^\perp \setminus \{\mathbf{0}\} : \mu_1(k_j) = d_j\}| \\
&\leq \sum_{\substack{d_1, \dots, d_s \in \mathbb{N}_0 \\ d_1 + \dots + d_s = z_1}} b^{d_1 + \dots + d_s - \rho_1(P) + 1} = b^{z_1 - \rho_1(P) + 1} \binom{z_1 + s - 1}{s - 1}.
\end{aligned}$$

Moreover, the bound on $|R_\alpha(\mathbf{k}, z_2)|$ obtained in Lemma 6 is independent of the choice of \mathbf{k} . Hence the result follows. \square

Finally we give a bound on $|J_\alpha(z)|$. Since $\mu_1(\mathbf{k}), \mu_1(\mathbf{l}) \geq \rho_1(P)$ for $\mathbf{k}, \mathbf{l} \in P^\perp \setminus \{\mathbf{0}\}$, $|J_\alpha(z)| = 0$ for $z < 2\rho_1(P)$. For $z \geq 2\rho_1(P)$ we have the following result.

Lemma 8. *Let P be an order $2\alpha + 1$ digital (t, m, s) -net over \mathbb{F}_b . For $z \geq 2\rho_1(P)$, we have*

$$|J_\alpha(z)| \ll_{\alpha, b, s, t} (z - 2\rho_1(P))^{2s\alpha + 1} z^{s-1} b^{(z - 2\rho_1(P))/2}.$$

Proof. For $z \geq 2\rho_1(P)$, we have

$$J_\alpha(z) = \bigcup_{\substack{z_1, z_2 \geq \rho_1(P) \\ z_1 + z_2 = z}} J_\alpha(z_1, z_2).$$

Since the sets on the right-hand side above are pairwise disjoint, we have

$$|J_\alpha(z)| = \sum_{\substack{z_1, z_2 \geq \rho_1(P) \\ z_1 + z_2 = z}} |J_\alpha(z_1, z_2)| \leq 2 \sum_{\substack{\rho_1(P) \leq z_1 \leq \lfloor z/2 \rfloor \\ z_2 = z - z_1}} |J_\alpha(z_1, z_2)|$$

where the last inequality stems from the symmetry of the Walsh coefficients, i.e., $\hat{K}_{\alpha, s}(\mathbf{l}, \mathbf{k}) = \hat{K}_{\alpha, s}(\mathbf{k}, \mathbf{l})$ for any $\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s$. By using the bound on $|J_\alpha(z_1, z_2)|$ given in Lemma 7, we have

$$|J_\alpha(z)| \ll_{\alpha, b, s} \sum_{\substack{\rho_1(P) \leq z_1 \leq \lfloor z/2 \rfloor \\ z_2 = z - z_1}} b^{z_1 - \rho_1(P) + 1} \binom{z_1 + s - 1}{s - 1} \prod_{i=1}^{2\alpha} \binom{(i+1)z_2 - \rho_{i+1}(P) + s}{s}.$$

It follows from Lemmas 2 and 3 that for $1 \leq i \leq 2\alpha$

$$\begin{aligned}
\binom{(i+1)z_2 - \rho_{i+1}(P) + s}{s} &\leq \binom{(i+1)(z_2 - m) + \lceil t(i+1)/(2\alpha+1) \rceil + s - 1}{s} \\
&\ll_{\alpha, b, s, t} (z_2 - \rho_1(P))^s.
\end{aligned}$$

Also we have

$$\binom{z_1 + s - 1}{s - 1} \ll_s z_1^{s-1}.$$

Hence

$$\begin{aligned} |J_\alpha(z)| &\ll_{\alpha,b,s,t} \sum_{\substack{\rho_1(P) \leq z_1 \leq \lfloor z/2 \rfloor \\ z_2 = z - z_1}} b^{z_1 - \rho_1(P) + 1} z_1^{s-1} (z_2 - \rho_1(P))^{2\alpha s} \\ &\ll_{b,s} (z - 2\rho_1(P)) b^{(z - 2\rho_1(P))/2} z^{s-1} (z - 2\rho_1(P))^{2\alpha s}. \end{aligned}$$

Thus the result follows. \square

4 Summability of the Walsh coefficients

In this section we prove:

Lemma 9. For $\alpha, s \in \mathbb{N}$,

$$\sum_{\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s} |\hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l})| < \infty.$$

We would recall that this summability is needed to ensure the pointwise absolute convergence of the Walsh series of $K_{\alpha,s}$.

Proof. For $\alpha \geq 2$, by using (1) and (2), we have

$$\begin{aligned} \sum_{\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s} |\hat{K}_{\alpha,s}(\mathbf{k}, \mathbf{l})| &= \sum_{\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s} \prod_{j=1}^s |\hat{K}_{\alpha,s}(k_j, l_j)| = \left(\sum_{k, l \in \mathbb{N}_0} |\hat{K}_\alpha(k, l)| \right)^s \\ &\leq D_{\alpha,b}^s \left(\sum_{k, l \in \mathbb{N}_0} b^{-\mu_\alpha(k) - \mu_\alpha(l)} \right)^s = D_{\alpha,b}^s \left(\sum_{k \in \mathbb{N}_0} b^{-\mu_\alpha(k)} \right)^{2s}. \end{aligned}$$

As shown in [5, Lemma 2.10], the last sum is bounded above by $1 + \alpha + b^{-2}$. Thus the result follows.

Let $\alpha = 1$. Since the sum $\sum_{k \in \mathbb{N}_0} b^{-\mu_1(k)}$ is no longer finite, we need to exploit the sparsity of the Walsh coefficients obtained in Proposition 1. Since

$$\sum_{\mathbf{k}, \mathbf{l} \in \mathbb{N}_0^s} |\hat{K}_{1,s}(\mathbf{k}, \mathbf{l})| = \left(\sum_{k, l \in \mathbb{N}_0} |\hat{K}_1(k, l)| \right)^s \leq D_{1,b}^s \left(\sum_{\substack{k, l \in \mathbb{N}_0 \\ \hat{K}_1(k, l) \neq 0}} b^{-\mu_1(k) - \mu_1(l)} \right)^s,$$

it suffices to prove that the last sum is finite. It follows from Proposition 1 that $\hat{K}_1(k, l) = 0$ whenever (k, l) is of type (p, q) with $p + q > 2$, so that

$$\begin{aligned} \sum_{\substack{k, l \in \mathbb{N}_0 \\ \hat{K}_1(k, l) \neq 0}} b^{-\mu_1(k) - \mu_1(l)} &= \sum_{\substack{p, q \in \mathbb{N}_0 \\ p+q \leq 2}} \sum_{\substack{k, l \in \mathbb{N}_0 \\ (k, l): \text{ type } (p, q)}} b^{-\mu_1(k) - \mu_1(l)} \\ &= \sum_{\substack{p, q \in \mathbb{N}_0 \\ p+q \leq 2}} \sum_{z_1, z_2 \in \mathbb{N}_0} b^{-z_1 - z_2} |\mathcal{J}_{p,q}(z_1, z_2)|, \end{aligned} \quad (11)$$

where we define the set

$$\mathcal{J}_{p,q}(z_1, z_2) := \{(k, l) \in \mathbb{N}_0^s : \mu_1(k) = z_1, \mu_1(l) = z_2, (k, l) \text{ is of type } (p, q)\}.$$

In what follows, we estimate the cardinality of $\mathcal{J}_{p,q}(z_1, z_2)$ and the inner sum of (11) for possible choices of p and q . Due to the symmetry of k and l , we only consider the case $p \geq q$.

1. $p = q = 0$: Since it follows that $k = l$, $\mathcal{J}_{p,q}(z_1, z_2)$ is empty if $z_1 \neq z_2$. Otherwise, the cardinality of $\mathcal{J}_{p,q}(z_1, z_2)$ equals the number of possible choices for k such that $\mu_1(k) = z_1$. Thus we have

$$|\mathcal{J}_{p,q}(z_1, z_2)| = \begin{cases} 0 & \text{if } z_1 \neq z_2, \\ \lceil b^{z_1-1}(b-1) \rceil & \text{otherwise,} \end{cases}$$

and

$$\sum_{z_1, z_2 \in \mathbb{N}_0} b^{-z_1-z_2} |\mathcal{J}_{p,q}(z_1, z_2)| = \sum_{z_1 \in \mathbb{N}_0} b^{-2z_1} \lceil b^{z_1-1}(b-1) \rceil = 1 + \frac{1}{b}.$$

2. $p = 1, q = 0$: Since k must be of the form $\kappa_1 b^{z_1-1} + l$, $\mathcal{J}_{p,q}(z_1, z_2)$ is empty if $z_1 \leq z_2$. Otherwise, the cardinality of $\mathcal{J}_{p,q}(z_1, z_2)$ equals the number of possible choices for l such that $\mu_1(l) = z_2$ times the number of possible choices for $\kappa_1 \in \{1, \dots, b-1\}$. Thus we have

$$|\mathcal{J}_{p,q}(z_1, z_2)| = \begin{cases} 0 & \text{if } z_1 \leq z_2, \\ \lceil b^{z_2-1}(b-1) \rceil (b-1) & \text{otherwise,} \end{cases}$$

and

$$\begin{aligned} \sum_{z_1, z_2 \in \mathbb{N}_0} b^{-z_1-z_2} |\mathcal{J}_{p,q}(z_1, z_2)| &= \sum_{z_1 > z_2} b^{-z_1-z_2} \lceil b^{z_2-1}(b-1) \rceil (b-1) \\ &\leq (b-1) \sum_{z_1 > z_2} b^{-z_1-z_2} b^{z_2} \\ &= (b-1) \sum_{z_1=1}^{\infty} b^{-z_1} \sum_{z_2=0}^{z_1-1} 1 \\ &= (b-1) \sum_{z_1=1}^{\infty} b^{-z_1} z_1 < \infty. \end{aligned}$$

3. $p = 2, q = 0$: Since k must be of the form $\kappa_1 b^{z_1-1} + \kappa_2 b^{c_2-1} + l$, $\mathcal{J}_{p,q}(z_1, z_2)$ is empty if $z_1 \leq z_2 + 1$. Otherwise, the cardinality of $\mathcal{J}_{p,q}(z_1, z_2)$ equals the number of possible choices for l such that $\mu_1(l) = z_2$ times the number of possible choices for $\kappa_1, \kappa_2 \in \{1, \dots, b-1\}$ and $c_2 \in \{z_2+1, \dots, z_1-1\}$. Thus we have

$$|\mathcal{J}_{p,q}(z_1, z_2)| = \begin{cases} 0 & \text{if } z_1 \leq z_2 + 1, \\ \lceil b^{z_2-1}(b-1) \rceil (b-1)^2 (z_1 - z_2 - 1) & \text{otherwise,} \end{cases}$$

and

$$\sum_{z_1, z_2 \in \mathbb{N}_0} b^{-z_1-z_2} |\mathcal{J}_{p,q}(z_1, z_2)|$$

$$\begin{aligned}
&= \sum_{z_1 > z_2 + 1} b^{-z_1 - z_2} \lceil b^{z_2 - 1} (b - 1) \rceil (b - 1)^2 (z_1 - z_2 - 1) \\
&\leq (b - 1)^2 \sum_{z_1 > z_2 + 1} b^{-z_1 - z_2} b^{z_2} (z_1 - z_2 - 1) \\
&\leq (b - 1)^2 \sum_{z_1 = 2}^{\infty} b^{-z_1} \sum_{z_2 = 0}^{z_1 - 2} (z_1 - z_2 - 1) \\
&= \frac{(b - 1)^2}{2} \sum_{z_1 = 2}^{\infty} b^{-z_1} (z_1 - 1) z_1 < \infty.
\end{aligned}$$

4. $p = q = 1$: Since k and l are given of the forms $k = \kappa_1 b^{z_1 - 1} + k'$ and $l = \lambda_1 b^{z_2 - 1} + k'$, respectively, with some $0 \leq k' < b^{\min(z_1, z_2) - 1}$ and $\kappa_1 b^{z_1 - 1} \neq \lambda_1 b^{z_2 - 1}$, the cardinality of $\mathcal{J}_{p,q}(z_1, z_2)$ equals the number of possible choices for k' times the number of possible choices for $\kappa_1, \lambda_1 \in \{1, \dots, b - 1\}$. Note that if $z_1 = z_2$, we have $\kappa_1 \neq \lambda_1$. Thus we have

$$\begin{aligned}
|\mathcal{J}_{p,q}(z_1, z_2)| &= \begin{cases} \lceil b^{\min(z_1, z_2) - 1} \rceil (b - 1)(b - 2) & \text{if } z_1 = z_2, \\ \lceil b^{\min(z_1, z_2) - 1} \rceil (b - 1)^2 & \text{otherwise,} \end{cases} \\
&\leq b^{\min(z_1, z_2)} (b - 1),
\end{aligned}$$

and

$$\begin{aligned}
\sum_{z_1, z_2 \in \mathbb{N}_0} b^{-z_1 - z_2} |\mathcal{J}_{p,q}(z_1, z_2)| &= (b - 1) \sum_{z_1, z_2 \in \mathbb{N}_0} b^{-z_1 - z_2} b^{\min(z_1, z_2)} \\
&\leq 2(b - 1) \sum_{z_1 \geq z_2 \geq 0} b^{-z_1 - z_2} b^{z_2} \\
&= 2(b - 1) \sum_{z_1 = 0}^{\infty} b^{-z_1} \sum_{z_2 = 0}^{z_1} 1 \\
&= 2(b - 1) \sum_{z_1 = 0}^{\infty} b^{-z_1} (z_1 + 1) < \infty.
\end{aligned}$$

In this way, since we only have six choices for p and q , and for each choice the inner sum of (11) is shown to be finite, (11) itself is also finite. Thus the result follows. \square

Acknowledgement

The work of T. G. is supported by JSPS Grant-in-Aid for Young Scientists No.15K20964. The work of K. S. and T. Y. is supported by Australian Research Council's Discovery Projects funding scheme (project number DP150101770). The authors are grateful to the reviewers for many helpful comments, particularly, for asking for the endpoint case $\alpha = 1$.

References

- [1] N. Aronszajn, Theory of reproducing kernels, Trans. Amer. Math. Soc. 68 (1950), 337–404.

- [2] J. Baldeaux and J. Dick, QMC rules of arbitrary high order: Reproducing kernel Hilbert space approach, *Constr. Approx.* 30 (2009), 495–527.
- [3] J. Baldeaux, J. Dick and F. Pillichshammer, Duality theory and propagation rules for higher order nets, *Discrete Math.* 311 (2011), 362–386.
- [4] W. W. L. Chen and M. M. Skriganov, Explicit constructions in the classical mean squares problem in irregularities of point distribution, *J. Reine Angew. Math.* 545 (2002), 67–95.
- [5] J. Dick, Explicit constructions of quasi-Monte Carlo rules for the numerical integration of high-dimensional periodic functions, *SIAM J. Numer. Anal.* 45 (2007), 2141–2176.
- [6] J. Dick, Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order, *SIAM J. Numer. Anal.* 46 (2008), 1519–1553.
- [7] J. Dick, The decay of the Walsh coefficients of smooth functions, *Bull. Aust. Math. Soc.* 80 (2009), 430–453.
- [8] J. Dick, Higher order scrambled digital nets achieve the optimal rate of the root mean square error for smooth integrands, *Ann. Statist.* 39 (2011), 1372–1398.
- [9] J. Dick, F. Y. Kuo, Q. T. Le Gia, D. Nuyens and C. Schwab, Higher order QMC Petrov–Galerkin discretization for affine parametric operator equations with random field inputs, *SIAM J. Numer. Anal.* 52 (2014), 2676–2702.
- [10] J. Dick, Q. T. Le Gia and C. Schwab, Higher order quasi-Monte Carlo integration for holomorphic, parametric operator equations, *SIAM/ASA J. Uncertainty Quantification* 4 (2016), 48–79.
- [11] J. Dick and F. Pillichshammer, *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, Cambridge, 2010.
- [12] J. Dick and F. Pillichshammer, Optimal \mathcal{L}_2 discrepancy bounds for higher order digital sequences over the finite field \mathbb{F}_2 , *Acta Arith.* 162 (2014), 65–99.
- [13] H. Faure, Discrepances de suites associées à un système de numération (en dimension s), *Acta Arith.* 41 (1982), 337–351.
- [14] K. K. Frolov, Upper error bounds for quadrature formulas on function classes, *Dokl. Akad. Nauk SSSR* 231 (1976), 818–821.
- [15] T. Goda, K. Suzuki and T. Yoshiki, Optimal order quasi-Monte Carlo integration in weighted Sobolev spaces of arbitrary smoothness, *IMA J. Numer. Anal.* (2016) DOI:10.1093/imanum/drw011.
- [16] T. Goda, K. Suzuki and T. Yoshiki, An explicit construction of optimal order quasi-Monte Carlo rules for smooth integrands, *SIAM J. Numer. Anal.* 54 (2016), 2664–2683.

- [17] J. H. Halton, On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numer. Math.* 2 (1960), 84–90.
- [18] F. J. Hickernell, A generalized discrepancy and quadrature error bound, *Math. Comp.* 67 (1998), 299–322.
- [19] F. J. Hickernell, Obtaining $O(N^{-2+\epsilon})$ convergence for lattice quadrature rules, In: K. T. Fang, F. J. Hickernell and H. Niederreiter, eds., *Monte Carlo and Quasi-Monte Carlo Methods 2000*, Springer, Berlin, 2002, pp. 274–289.
- [20] F. J. Hickernell, My dream quadrature rule, *J. Complexity* 19 (2003), 420–427.
- [21] F. J. Hickernell, P. Kritzer, F. Y. Kuo and D. Nuyens, Weighted compound integration rules with higher order convergence for all N , *Numer. Algorithms* 59 (2012), 161–183.
- [22] A. Hinrichs, L. Markhasin, J. Oettershagen and T. Ullrich, Optimal quasi-Monte Carlo rules on order 2 digital nets for the numerical integration of multivariate periodic functions, *Numer. Math.* 134 (2016), 163–196.
- [23] D. Kreig and E. Novak, A universal algorithm for multivariate integration, *Found. Comput. Math.* (2016) DOI:10.1007/s10208-016-9307-y.
- [24] F. Y. Kuo, I. H. Sloan and H. Woźniakowski, Periodization strategy may fail in high dimensions, *Numer. Algorithms* 46 (2007), 369–391.
- [25] V. K. Nguyen, M. Ullrich and T. Ullrich, Change of variable in spaces of mixed smoothness and numerical integration of multivariate functions on the unit cube, *ArXiv Preprint arXiv:1511.02036*.
- [26] H. Niederreiter, Low-discrepancy point sets, *Monatsh. Math.* 102 (1986), 155–167.
- [27] H. Niederreiter, Low-discrepancy and low-dispersion sequences, *J. Number Theory* 30 (1988), 51–70.
- [28] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 63, SIAM, Philadelphia, 1992.
- [29] H. Niederreiter and C. P. Xing, *Rational Points on Curves over Finite Fields: Theory and Applications*, London Mathematical Society Lecture Note Series 285, Cambridge University Press, Cambridge, 2001.
- [30] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems, Volume I: Linear Information*, EMC Tracts in Mathematics 6, European Mathematical Society, Zürich, 2008.
- [31] A. B. Owen, A constraint on extensible quadrature rules, *Numer. Math.* 132 (2016), 511–518.
- [32] M. Yu. Rosenbloom and M. A. Tsfasman, Codes for the m -metric, *Probl. Inf. Transm.* 33 (1997), 55–63.

- [33] M. M. Skriganov, Harmonic analysis on totally disconnected groups and irregularities of point distributions, *J. Reine Angew. Math.* 600 (2006), 25–49.
- [34] I. M. Sobol', The distribution of points in a cube and approximate evaluation of integrals, *Zh. Vychisl. Mat. i Mat. Fiz.* 7 (1967), 784–802.
- [35] S. Tezuka, Polynomial arithmetic analogue of Halton sequences, *ACM Trans. Model. Comput. Simul.* 3 (1993), 99–107.
- [36] M. Ullrich, On “Upper error bounds for quadrature formulas on function classes” by K. K. Frolov, In: R. Cools and D. Nuyens, eds., *Monte Carlo and Quasi-Monte Carlo Methods 2014*, Springer, Heidelberg, 2016, pp. 571–582.
- [37] M. Ullrich and T. Ullrich, The role of Frolov’s cubature formula for functions with bounded mixed derivative, *SIAM J. Numer. Anal.* 54 (2016), 969–993.
- [38] G. Wahba, *Spline Models for Observational Data*, CBMS-NSF Regional Conference Series in Applied Mathematics 59, SIAM, Philadelphia, 1990.